

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:17 ; Search time 299.73 Seconds
(without alignments)
244.403 Million cell updates/sec

Title: US-09-331-631A-5

Perfect score: 3326
Sequence: 1 OCMOLETSGQMRKVCSCDK.....SPRSTKQOQPLVSLIDFVGF 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:archaea:*
2: SP:bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_minc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unklassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3326	100.0	625	10	Q9SP13 macadamia i
2	3327	97.0	666	10	Q9SP15 macadamia i
3	3215	96.7	666	10	Q9SP14 macadamia i
4	1310.5	39.4	593	10	Q9SEW4 juglans reg
5	1163	35.0	810	10	Q9ZWI3 cucurbita m
6	1015	30.5	525	10	Q43358 theobroma c
7	888	26.7	582	10	Q03865 zea mays (m
8	864.5	26.0	637	10	Q03678 hordeum vul
9	844.5	25.4	613	10	Q9M3X6 pisum sativ
10	827	24.9	544	10	Q22120 glycine max
11	811	24.4	545	10	Q41674 vicia narbo
12	777.5	23.4	489	10	Q9SP11 glycine max
13	775.5	23.3	483	10	Q49927 pisum sativ
14	758	22.8	417	10	Q22121 glycine max
15	757.5	22.8	438	10	Q43626 pisum sativ
16	749.5	22.5	463	10	Q41677 vicia narbo
17	747	22.5	486	10	Q91UJ7 arabidopsis
18	740	22.2	448	10	Q40873 picea glauc
19	729	21.9	450	10	Q40844 picea glauc

20	709.5	21.3	518	10	Q9M3X8 lens culina
21	649	19.5	461	10	Q9ZRG9 oryza sativ
22	560	16.8	430	10	Q41115 phaseolus v
23	558.5	16.8	414	10	Q41727 zania furfu
24	556	16.7	430	10	Q43633 phaseolus v
25	549	16.5	421	10	Q43632 phaseolus v
26	541.5	16.3	232	10	Q9SQ49 rullingia ma
27	532.5	16.0	540	10	Q03866 zea mays (m
28	527.5	15.9	423	10	Q43617 phaseolus l
29	527	15.8	229	10	Q9SQ50 abroma angu
30	522	15.7	239	10	Q9SQ42 theobroma b
31	521.5	15.7	236	10	Q9SQ41 theobroma c
32	521.5	15.7	428	10	Q40913 phaseolus l
33	521	15.7	238	10	Q9SQ32 theobroma s
34	520.5	15.6	236	10	Q9SQ47 theobroma s
35	520	15.6	239	10	Q9SQ44 theobroma v
36	518.5	15.6	236	10	Q9SQ40 theobroma c
37	517.5	15.6	236	10	Q9SQ36 theobroma g
38	517.5	15.6	236	10	Q9SQ34 theobroma m
39	516.5	15.5	236	10	Q9SQ37 theobroma g
40	516	15.5	239	10	Q9SQ35 theobroma c
41	516	15.5	239	10	Q9S7V9 theobroma c
42	515.5	15.5	236	10	Q9SQ48 theobroma m
43	515	15.5	239	10	Q9SQ43 theobroma b
44	510.5	15.3	236	10	Q9SQ33 theobroma m
45	508	15.3	235	10	Q9SQ38 theobroma g

ALIGNMENTS

RESULT 1
ID Q9SP13 PRELIMINARY: PRT: 625 AA.
AC Q9SP13;
DT 01-MAY-2000 (TRENBLREL, 13, Created)
DT 01-OCT-2000 (TRENBLREL, 13, last sequence update)
DE VICILIN PRECURSOR (FRAGMENT).
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=606598;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NOT KNOWN.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia.";
RL Plant J. 0:0-0(1999).
DR EMBL: AF161885; MADS4246.1; -.
DR HSP: P02853; 2PIL.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
FT NON_TER
SQ SEQUENCE 625 AA: 73586 MW: 415808489D370296 CRC64:

Query Match 100.0%; Score 3326; DB 10; Length 625;
Best Local Similarity 100.0%; Pred. No. 7.7e-236;
Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCMOLETSGQMRKVCSCDKREFEDIDWSKYNQEDPOTECOCQRCRQOQSDPPOQY 60
DB 1 OCMOLETSGQMRKVCSCDKREFEDIDWSKYNQEDPOTECOCQRCRQOQSDPPOQY 60
QY 61 CORRKEICEEEBEYNNRODPOQYEQCCRCORRETEPRHMOICQRCRRERYEKKRQ 120
DB 61 CORRKEICEEEBEYNNRODPOQYEQCCRCORRETEPRHMOICQRCRRERYEKKRQ 120
QY 121 ORRYEORQREDEKYEKKRQEDDKRRDPOQYEQCCRCRRCEQOEPRLOYOCRCRQOQDOR 180
DB 121 ORRYEORQREDEKYEKKRQEDDKRRDPOQYEQCCRCRRCEQOEPRLOYOCRCRQOQDOR 180

```
Db 121 QKRYEQDEDEEYKYEERKKEGDNKRDPQOREYEDCRRHCEQOEPRLOYOCORRCQEQOR 180
QY 181 QHGRGDDLMPORGSGRGREEGEEKOSDNPPYFEDERSLSTRTEREGHISVLENYGRSK 240
Db 181 QHGRGDDLMPORGSGRGREEGEEKOSDNPPYFEDERSLSTRTEREGHISVLENYGRSK 240
QY 241 LRLALKNYRLVLEANPNFVLPHTLDADAILLVIGRGALKMIRHNDRESYNLECGDYI 300
Db 241 LRLALKNYRLVLEANPNFVLPHTLDADAILLVIGRGALKMIRHNDRESYNLECGDYI 300
QY 301 RLPAGTFFYLNRDNNERLHIAKFLQTTSTPGQYKEFFPAGGONPEPYLSTFSKELLEA 360
Db 301 RLPAGTFFYLNRDNNERLHIAKFLQTTSTPGQYKEFFPAGGONPEPYLSTFSKELLEA 360
QY 361 LNTOTERLGVLGQOQREGYIIRASQEQIRELTREDSESRHWHIRRGESSRGYPYLPFKR 420
Db 361 LNTOTERLGVLGQOQREGYIIRASQEQIRELTREDSESRHWHIRRGESSRGYPYLPFKR 420
QY 421 PLYSNKYGAYEYKPEPDYRQLODMDSVFITANTOGSMGPFENRSTKVVVVVAGSEADY 480
Db 421 PLYSNKYGAYEYKPEPDYRQLODMDSVFITANTOGSMGPFENRSTKVVVVVAGSEADY 480
QY 481 EMACPHLSGRHGGRGGKRRHEEEYHYEQVBARLSKREAIIVLAGHPVYVSSGNENLL 540
Db 481 EMACPHLSGRHGGRGGKRRHEEEYHYEQVBARLSKREAIIVLAGHPVYVSSGNENLL 540
QY 541 LFAFGINAONNHNENFLAGERNVLOOIEPOMELAFASRKEVEELFNSODESIFPPGPR 600
Db 541 LFAFGINAONNHNENFLAGERNVLOOIEPOMELAFASRKEVEELFNSODESIFPPGPR 600
QY 601 QHQOQSPRSTKQOQPLVSTLDFVGF 625
Db 601 QHQOQSPRSTKQOQPLVSTLDFVGF 625

RESULT 2
QSP15 PRELIMINARY: PRT: 666 AA.
AC QSP15:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL;
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia kernels.";
RL Plant J. 0:0-0(1999)
DR EMBL: AF161883; AAD54244.1; -
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA: 78217 MW: C752B884B2DP0224 CRC64;
```

Query Match 97.0%; Score 3227; DB 10; Length 666;
Best local Similarity 96.6%; Pred. No. 1.5e-228;
Matches 604; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

```
QY 1 OCMQLETSGOMRCVSCQCKRFEEDIDMSKYDNOEDPOTECOCORRCQOESDPQOQY 60
Db 42 OCMQLETSGOMRCVSCQCKRFEEDIDMSKYDNOEDPOTECOCORRCQOESDPQOQY 101
QY 61 CORCKEICEEEYENRORPQOQYECOCRCORRETEPHIMOICORCERRYEXEKRXO 120
Db 102 CORCKEICEEEYENRORPQOQYECOCRCORRETEPHIMOICORCERRYEXEKRXO 161
```

```
QY 121 QKRYEQDEDEEYKYEERKKEGDNKRDPQOREYEDCRRHCEQOEPRLOYOCORRCQEQOR 180
Db 162 QKRYEQDEDEEYKYEERKKEGDNKRDPQOREYEDCRRHCEQOEPRLOYOCORRCQEQOR 221
QY 181 QHGRGDDLMPORGSGRGREEGEEKOSDNPPYFEDERSLSTRTEREGHISVLENYGRSK 240
Db 222 QHGRGDDLMPORGSGRGREEGEEKOSDNPPYFEDERSLSTRTEREGHISVLENYGRSK 281
QY 241 LRLALKNYRLVLEANPNFVLPHTLDADAILLVIGRGALKMIRHNDRESYNLECGDYI 300
Db 282 LRLALKNYRLVLEANPNFVLPHTLDADAILLVIGRGALKMIRHNDRESYNLECGDYI 341
QY 301 RLPAGTFFYLNRDNNERLHIAKFLQTTSTPGQYKEFFPAGGONPEPYLSTFSKELLEA 360
Db 342 RLPAGTFFYLNRDNNERLHIAKFLQTTSTPGQYKEFFPAGGONPEPYLSTFSKELLEA 401
QY 361 LNTOTERLGVLGQOQREGYIIRASQEQIRELTREDSESRHWHIRRGESSRGYPYLPFKR 420
Db 402 LNTOTERLGVLGQOQREGYIIRASQEQIRELTREDSESRHWHIRRGESSRGYPYLPFKR 461
QY 421 PLYSNKYGAYEYKPEPDYRQLODMDSVFITANTOGSMGPFENRSTKVVVVVAGSEADY 480
Db 462 PLYSNKYGAYEYKPEPDYRQLODMDSVFITANTOGSMGPFENRSTKVVVVVAGSEADY 521
QY 481 EMACPHLSGRHGGRGGKRRHEEEYHYEQVBARLSKREAIIVLAGHPVYVSSGNENLL 540
Db 522 EMACPHLSGRHGGRGGKRRHEEEYHYEQVBARLSKREAIIVLAGHPVYVSSGNENLL 581
QY 541 LFAFGINAONNHNENFLAGERNVLOOIEPOMELAFASRKEVEELFNSODESIFPPGPR 600
Db 582 LFAFGINAONNHNENFLAGERNVLOOIEPOMELAFASRKEVEELFNSODESIFPPGPR 641
QY 601 QHQOQSPRSTKQOQPLVSTLDFVGF 625
Db 642 QHQOQSPRSTKQOQPLVSTLDFVGF 666
```

```
RESULT 3
QSP14 PRELIMINARY: PRT: 666 AA.
AC QSP14:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL;
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia.";
RL Plant J. 0:0-0(1999)
DR EMBL: AF161884; AAD54245.1; -
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA: 78243 MW: DECA22F8710F8A7B CRC64;
```

Query Match 96.7%; Score 3215; DB 10; Length 666;
Best local Similarity 96.6%; Pred. No. 1.2e-227;
Matches 604; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

```
QY 1 OCMQLETSGOMRCVSCQCKRFEEDIDMSKYDNOEDPOTECOCORRCQOESDPQOQY 60
Db 42 OCMQLETSGOMRCVSCQCKRFEEDIDMSKYDNOEDPOTECOCORRCQOESDPQOQY 101
```

QY	61	CORCKEICEEEREEENRPRDQOOOYEOOQKOCORRETEPHRPMOLCOORCRRREYKRRKQ	120
QY	102	CORCKEICEEEREEENRPRDQOOYEOOQECQNHETEPHMQOOCORCERRREYKRRKQ	151
QY	121	OKRYEEOQREDEEYEEERKMGDNKRDPOOREXEDCRHNCEDOEPRLOYOCORRCOOR	180
QY	162	OKRYEEOQREDEEYEEERKMGDNKRDPOOREYEDCKRRRCEDOEPRLOYOCORRCOOR	221
QY	181	OHGRGDDLMNPORGSSGYEEEGEEKOSDNPYFEDERSLSTRFTEEGHISVLENFYGRSK	240
QY	222	OHGRGDDLMNPORGSSGYEEEGEEKOSDNPYFEDERSLSTRFTEEGHISVLENFYGRSK	281
QY	241	LLRALKNRRLVLEAPNNAEYLPPLHLADAILVITGSGALKMIHRNDRRESYNLECGDVI	300
QY	282	LLRALKNRRLVLEAPNNAEYLPPLHLADAILVITGSGALKMIHRNDRRESYNLECGDVI	341
QY	301	RIPACTFTYLLINRONNEMRLIAKLOTITSPGOYKEFFPRAGGONPEBYLSTFSKEILLEA	360
QY	342	RIPACTFTYLLINRONNEMRLIAKLOTITSPGOYKEFFPRAGGONPEBYLSTFSKEILLEA	401
QY	361	LNTQGERLRGLVGGOOREVITIRASOEOIRELTRDSDSRMHIRRGESSRGYPNLFNKR	420
QY	402	LNTQGERLRGLVGGOOREVITIRASOEOIRELTRDSDSRMHIRRGESSRGYPNLFNKR	461
QY	421	PLYSNKYGOAYEVKPEDYROLQMDVSVFLANTTQSGMGPFFNTSTKVVVAASGADY	480
QY	462	PLYSNKYGOAYEVKPEDYROLQMDVSVFLANTTQSGMGPFFNTSTKVVVAASGADY	521
QY	481	EMACPRLSGRGGGGGGRHREEEEVYEOVRARLSKREAVILVLAGPVYFVSSGNLIL	540
QY	522	EMACPRLSGRGGGGGGRHREEEVDYVBOVKARLSKREAVILVPGHPVYFVSSGNLIL	581
QY	541	LFAFGINQNHNENFLAGRENNVLOQIEPOMELAFASKEVEEELFNSODESIFPPGPR	600
QY	582	LFAFGINQNHNENFLAGRENNVLOQIEPOMELAFAPKREVEEELFNSODESIFPPGPR	641
QY	601	OHOOOSPRSTKOOQPLVSIIDPVCF	625
QY	642	OHOOOSPRSTKOOQPLVSIIDPVCF	666
RESULT	4		
Q9SEMA	ID	Q9SEMA	PRT: 593 AA.
AC	Q9SEMA:	PRELIMINARY:	
DT	01-MAY-2000 (TREMblrel. 13, Created)		
DT	01-MAY-2000 (TREMblrel. 13, Last sequence update)		
DT	01-OCT-2000 (TREMblrel. 15, Last annotation update)		
DE	VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).		
OS	Juglans regia (English walnut).		
OC	Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:		
OC	Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eustosids I;		
OC	regales: Juglandaceae; Juglans.		
OX	NCBI_TaxID=51240;		
RN	11		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. SUNLAND. TISSUE=SOMATIC EMBRYO LINE:		
RA	Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;		
RT	"Identification and cloning of a cDNA encoding a vicillin-like protein		
RT	Jug r 2, from English walnut kernel (Juglans regia): a major food		
RT	allergen.";		
RL	Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.		
DR	EMBL: AF066055; AAF18269.1; -		
DR	HSSP: P02853; 2PHT.		
DR	INTERPRO: IPR001113; -		
DR	PRAM: PR00546; SeedsStore_7s; 1.		
FT	NON_TER		
QY	593 AA: 69990 MW: 9BA127E19B18C0D8 CRC64:		

Query Match	39.4%;	Score 1310.5;	DB 10;	Length 593;
Best Local Similarity	44.1%;	Pred. No. 4.8e-88;		
Matches 273;	Conservative 121;	Mismatches 166;	Indels 59;	Gaps 177;

OY	32	DNEODPTECCSOCCROEDPPOOYOCORKEITEE----	EEENPO-----	78
OY	32	DNEODPTECCSOCCROEDPPOOYOCORKEITEE----	EEENPO-----	78
OY	9	ENPRDPRREYKOCSDCKRGGOGGOGGOGGOCGICSEBRL	EEODRSDEENRRKRGDVBQD	68
OY	79	-RDPPOOYECOSCKORRETERPHNOICQORCERRYEK	RKREDOORDEDEKYE	13
OY	138	RKKEBDKRRDPOREYEDCRHNC--EODEPRLQYOCOR	SOEVO-QORNGR-GGDLMNFOR	19
OY	113	----GRDRODPQ--OYNRCORCKOIOECSPEOROC	OCORCEROYKEOGRGREASPRR	167
OY	194	GGSGYEGEGEKKOSNPPYFEDRSISTPRRNEGHIS	LVNMPYGRSKILRLKRYLVLL	25
OY	168	ESRGR---EEEOQRNPPYFHSQISIRSHNESEBE	GKYLEPTEKTELRLGIENTRVYL	222
OY	254	EANPAFVLPRLYDADALLVIGRGALKMTHRONRES	YNLECGDVJIRPAGTFYLLNR	31
OY	225	DANPNTSMRLPHNKAESVAUYTRGKATLTLVSOE	RESFNLECGDVLIRPAGATVYVINO	28
OY	314	DNNEPLTAKFLQTISTRGQYKEFFRPAQGNP-EP	YISTESKELLEALNTOTERLQVL	37
OY	285	DSNERLEWVKLLOPANNQCGREYVYAAKSPDOS	YLVEFNDLIVALNTPRDLREFF	344
OY	373	GO--ORGEVJIRASOEQIRBELTRD--DSESR	RNMTIRGCESSRQPNYLFNKRPLYSNKG	422
OY	345	DOOEDREGVJIRASOEKIRALSOMNASKGQRP	----GRSSGGISLKSSEPSISNFG	400
OY	429	QAYEVKREDYROLQDMDSVFLANTTQSGSMG	PEFFNPRSTVYVVVVASGEADAVENACPHLS	488
OY	401	QFEFACSRPHNLOEDMDYLVNVAELIKRGAMV	PHYNSKATVYVVVEBTGYEMACPHVS	460
OY	489	GRHGRGGGKRLEEBEY--HYEOYRALSKREAL	YVLAGHPVYVSSGNNLLFAFGI	544
OY	461	SO-SYEGSGRRROEBEESTGRFOKATRLAG	LDIFVLPAQHPRLATYSOMENLRLEFDI	519
OY	547	NAONNENFELGRENVLQOOLEROMALELAFAS	KREVELLESNODSEIFFPGRPHOQOS	600
OY	520	NGENNRQFLAG--QNNITNOLEREKELSF	MPREIETIETESOMESTFYPTER----	OS 574
OY	607	PRSTRQOQPLVSLDEYVGF	625	
OY	575	RRCQGRDHPPLASILDFAFF	593	
RESULT	5			
Q9ZW13	09%MI3	PRELIMINARY:	PRT: 810 AA.	
AC	09%MI3:			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	PV100.			
OS	Cucurbita maxima (Pumpkin) (Winter squash).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eustosids I;			
OC	Cucurbitales; Cucurbitaceae; Cucurbita.			
OX	NCBI_TaxID=3661:			
RP	11			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=KUROKAWA AMAKURI NANKIN; TISSUE=COTYLEDON;			
RC	MEDLINE=99107919; Pubmed=9891029;			
RA	Yamada K., Shimada T., Kondo M., Nishimura M., Hara-Nishimura I.;			
RT	"Multiple functional proteins are produced by cleaving Asn-Gln bonds			
RT	of a single precursor by vacuolar processing enzyme.";			
RL	J. Biol. Chem. 274:2563-2570(1999).			
DR	EMBL: AB019195; BAA34056.1; -			
DR	HSSP: P02853; 2PHL.			
DR	INTERPRO: IPR001113; -			
DR	PFAM: PF00546; Seedstore.7s; 1.			
DR	PRODOM: PD081059; -; 1.			
SO	SEQUENCE 810 AA; 97314 MW; A829A3F754226GAB CRC64;			

DR PFAM: PF00546: Seedstore_7s, 1.
DR PRODOM: PD081059; -, 1.
SO SEQUENCE 810 AA; 97314 MW; A829A3F7542266AB CXC64

Query Match 35.0%; Score 1163; DB 10; Length 810;
 Best local similarity 36.8%; Pred. No. 4,9e-77;
 Matches 271; Conservative 133; Mismatches 192; Indels 140; Gaps 22;

10 QMRVCSQCKRFE-----DIDMSKY----- 31
 94 OQRKCEOVEEERLEREOGEGEDVDVERDPEMEREOERREHERERERERERER 153
 32 -----DNOEDPOTEOCOCRCROOESDPROOQYOCRCKEICEPEEEENRORQOQ 85
 154 GCRRERENENDEPREREOEROR--REOERRREREOERERERGERDEDDEN--ORDPDMR 210
 86 EOCQRCQRRETERPRHMOICQORCERRY-----EKE-----KRCQKRYEEOQREDE 132
 211 EQ-----ERREOERRREREOERERORGRORGRORDEDEDNORDPDMRBOQRKREOERER 265
 133 EKEEEMKEDNKRDP-----OOREYEDCRRICEOO-----EP--R 166
 266 ERGGHDEDEDNORDPDMRREOERREOERERERERERERERERERERERERER 325
 167 LOYOCORCOEQ-----ORHGRG-----DLMPORGSGRYE-----GEEK 205
 326 REOERRREREOER 385
 206 QSDNPYFDESLSTRTEEGHISYLENFYGRSKLLRALKNYRLYLEANPNAFVLP 265
 386 QSNPPYFQEQRQSHRYRDEGHWRLERSEKSELLKIGNRLALERAPHTFPIH 445
 266 LDDALLIYIGGALKMHRDKR---ESYNECGDVIRIPAGTFYILNRNNERLH 321
 446 LDECVLLVGRATITTVQEKRETRKESYVESGDVMIIPAGTLLYLANOE--NEDLO 504
 322 AKFLQITSPGOKKEFFPAGONPEPYLSTFSKEILEAALNTQTERLGVLAGOORE--GV 379
 505 VKLVQVYNNPGEFKDYLISAGES--QAYISVFSNDVLEALNITRDKLERLFRKORER 563
 380 ITRASQEQIRELTRDSESRMRHIRGESSRPNMLFNKRPYSNKYGAVYEVKEDYR 439
 564 IYRASQEQIRALISQ-----RATSVKRGSGRVAPIKLESQTPVNNQYGMFACPD 618
 440 QLODMVSVFIANTOGSMNGPFFNTRSTKVVVAVASGEADVEMACHLSRHGRGGK- 498
 619 QURTVAVASVDIKQGGMMVPHFNSRAITVWVVSAGSFEFACPHISSQOMQRREE 678
 499 -----RHSEEE-----VHYEOVARLSKREAIYVLGHPVVFSSGNENLLPFGIN 549
 679 ERHMRREERERERERGRPRVAGRLSEGVLVYPAGHPLATMASRPNMLRVGFGIN 738
 550 NNHNFVLAGERNVLOQIEFOAMELAFASRKEVEELFNSQDESIFPGPRHOQOSPRS 609
 739 NNHNRNLAGE--NIMNELLDEAKELAFNVEGKQADELFRSQRESFTTEGP-----EGRRR 793
 610 TKOQOPLVSLDVG 625
 794 STERSPLSLTLKLAGY 809

RESULT 6
 043358 PRELIMINARY: PRT: 525 AA.

AC 043358;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE VICILIN PRECOURSOR.
 GN CSV.
 OS Theobroma cacao (Cacao).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Malvales; Malvaceae; Theobroma.
 NCBI_TaxID=3641;

RN 111
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAVES;
 RX MEDLINE=92288309; PubMed=1600151;
 RA McHenry L., Fritz P.J.,
 RT "Comparison of the structure and nucleotide sequences of vicilin genes
 of cocoa and cotton raise questions about vicilin evolution.";
 RL Plant Mol. Biol. 18:1173-1176(1992).
 DR EMBL: X62625; CAA44493.1; -;
 DR EMBL: X62626; CAA44494.1; -;
 DR HSSP: P02853; 2PHL.
 DR MENDEL: 30919; TheCC:1188; 30919.
 DR INTERPRO: IPR001113; -;
 DR PFAM: PF00546; Seedstore_7s; 1.
 DR PRODOM: PD081059; -; 1.
 KW SIGNAL.
 FT SIGNAL. 1 24 POTENTIAL.
 FT CHAIN 25 525 VICILIN.
 SQ SEQUENCE 525 AA; 60798 MW; 19114CD5C248905D CRC64;

Query Match 30.5%; Score 1015; DB 10; Length 525;
 Best local similarity 41.1%; Pred. No. 1,9e-66;
 Matches 211; Conservative 91; Mismatches 158; Indels 54; Gaps 11;

68 ICEEEELYNR--ORDPQOYEOCQRCQRRETERPRHMOICQORCERYEKEKQOKRY 124
 22 LCGSAYAGRKQYERDPPOQYEOCQRCSEATEEREOECQRCR-----EX 70
 125 EEOQREDEEYEREMKEGNKRPDOOREYEDCRRHCEOE--PRLOYOCORCOEOROH 182
 71 KEQOROEEL-----QROYOCCGRCEODQOOGREOOCQRCMEYKQO 116
 183 GRGGLMNPORGSGRYEEGEEKSDNPYFDE--RSLSTRFTEEGHISYLENFYGRSKL 241
 117 ER--GEHENYHNHKKNSSEEEGQKNNPYFRKRSFQFRDEEGNFKLQFAENSP 175
 242 LRALKNYRLYLEANPNAFVLPHTLADATLLVIGRGALKMHRDNRESYNLECGDVIR 301
 176 LKGINRYLAMEFANNTFTLPHHCDAEALYFTNGKGTTFVTHNKESYNVQRGTYS 235
 302 IPAQTFYILNRNNRNLHAKFLQITSPGOKKEFFPAGONPEPYLSTFSKEILEAL 361
 236 VPAGSTVYVSQONKRLTAVLALPVNSPGKYLEFPAGNNPEYVYAGFVYLETYF 295
 362 NTQTERLGVLAGOOR-----EGYIIRASQEQIRELTRDSESRMRHIRGESSRGP 413
 296 NTQREKLEILEOROGKROGQGGMFRAKAPQIRALISQATSPR-----HRGE--RLA 349
 414 YNLNRPYLSNKYGAVYEVKEDYRQLODMVSVFIANTOGSMNGPFFNTRSTKVVV 473
 350 INLLSQSPVSNQGRFFECAPEDFSQFMQDVAVSAFLNQCAIFVPHNSKATVFEV 409
 474 ASGEADVEMACHLSRHGRGGK--RHSEEEV-----HYEOVARLSKREAIYVLA 525
 410 TDQGYAQMACHPLSRQOSQSGQRODRROEESSEETFGFQYQKAPLSPGDVEVABA 469
 526 GHPVVFSSGNENLLPFGINQNNHNFVLAG 559
 470 GHAVTFPASKDPLNAVAFGLNQNNOIRIFLAG 503

RESULT 7
 003865 PRELIMINARY: PRT: 582 AA.

AC 003865;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE VICILIN-LIKE EMBRYO STORAGE PROTEIN.
 GN GLB1-L.
 OS Zea mays (maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
OX NCBI_TaxID=4577;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=M64A;
RX MEDLINE=92090707; PubMed=1752424;
RA Belanger F.C., Kitz A.L.;
RT "Molecular basis for allelic polymorphism of the maize globulin-1
gene."
RL Genetics 129:863-872(1991).
R EMBL: X59083; CAA1809.1; -.
DR HSSP: P50477; ICAN.
DR MENDEL: 11234; Zeama:1188;11234.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
DR PROSITE: PS00867; CPASASE_2; UNKNOWN_1.
DR PRODOM: PD081059; -; 1.
KW Seed storage protein.
SQ SEQUENCE 582 AA; 66162 MW; 680D85FEC07C885 CRC64;

Query Match 26.7%; Score 888; DB 10; Length 582;

Best Local Similarity 36.7%; Pred. No. 4.5e-57;
Matches 208; Conservative 87; Mismatches 181; Indels 90; Gaps 13;

OY 73 EEYNRORDPQOQYEOQRCORCERETPRHMQ-ICQORCERHYEKERKQOQRYEEQOQRED 131
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 25 EDNNHHNGHNGKSGOCYRCEDR--RWNRPRKLEOC-REERERKQERSRHEADDRSG 80
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 132 EEYEEYEMKGDNRKDRQOREYEDSRNSECQERLQYOCQRRQOQOQNRGSGDLMNP 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 81 EGSSSD-----FREDE----- 91
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 192 QRGSGRGEYEGEKOSD-NPYUFEDERSLSTRFTEEGHISYLENFYGRSKLLRAKLNRYL 250
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 92 -----KQKQDRPRYVDRSRFRVRSYEOGSLVLRPEVSLRLGLGINDYRV 140
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 251 VLEANDPNAFVLPYTHLDADAILLVIGRGALKMIRHNDRESYNLECGDVIIRIPAGTTFYL 310
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 141 AVELEANDPNSVYVSHNDACICVAGCEGVVTTIENGERSYTIKQGNFVARAGAVTYL 200
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 311 INDDNNERLIAKFLQITSTPGQYKEFPFAGGONPERPYLSTESKELLEALNTQTERLKG 370
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 201 ANPDGKKKLVITKILHTISVPGFQFFGCGGRNPESFLSSFESKTQRAAYKTSQRLER 260
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 371 VLGO--OREGVIIIRASQEOIRLETRDSE---SRHWHIRGSGSSGPYNLFKRPRLYSN 425
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 261 LGRHGDQDKGILYKATEEOTRELRRHASGEGHNPRLPRFG-SKGPYSLDRPSIAN 319
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 426 KYQAAVEVPREDYROLQMDVSVFIANTITOGSMGPFNTSTKVVVVASGEADVEMACP 485
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 320 OHQQLLEADARSPHDLAEHDVSVFANITAGSMARLFTMRSPKIAVYVNGKQYATIVCP 379
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 486 HLSGRIG---GRGGKRHEEEVEVH-----YEOVRARLSKREAIIVLAGHPVVF 531
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 380 HROSQSGESEREEDKGRREEESESESESESESESESESESESESESESESESESESESE 439
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 532 VSSGNNLLLFAGGINAONNHENFLAGRENVLDQTEPQAMELAPASKEVEBELNSOD 591
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 440 VASRDNLQIVCEVYADNRKEKFLAGAD-NVLOKLDRAKALSTFASKAEVDEVGLSKRR 498
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 592 ESTFPFPRPQ---HOOSPRSTKQOQ 614
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 499 EKGFLPGPESGGHEERQDEEREE 524
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
RESULT 8
003678 PRELIMINARY; PRT: 637 AA.
ID 003678
AC 003678
DT 01-NOV-1996 (TREMBlrel. 01, Created)
01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE FETAL STORAGE PROTEIN.
GN BEG1 OR GBL1.
OS Hordeum vulgare (Barley), and Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
OX NCBI_TaxID=4513, 4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93287988; PubMed=8510647;
RA Heck G.R., Chamberlain A.C., Ho T.H.D.;
RT "Barley embryo globulin 1 gene, Beg1: characterization of cDNA,
chromosome mapping and regulation of expression."
RL Mol. Gen. Genet. 239:209-218(1993).
DR EMBL: M64372; AAA32936.1; -.
DR EMBL: M81719; AAA34269.1; -.
DR HSSP: P02853; 2PHL.
DR MENDEL: 8553; Horvu:1188;8553.
DR INTERPRO: IPR000901; -.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
DR PROSITE: PS00867; CPASASE_2; UNKNOWN_1.
DR PRODOM: PD081059; -; 1.
KW Seed storage protein.
SQ SEQUENCE 637 AA; 72252 MW; F323FAFF99947C3C CRC64;

Query Match 26.0%; Score 864.5; DB 10; Length 637;

Best Local Similarity 35.0%; Pred. No. 2.7e-55;
Matches 221; Conservative 100; Mismatches 219; Indels 91; Gaps 19;

OY 29 SKYDNEBPQ--TECOQCORCQODESDPQOQYCCORCKEICEEEBEYNRDPQOQYE 86
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 27 ASHDEDDDRGGHSIQOCYORCQER--PR---YSNARKVQEC-----RDDQOH- 71
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 87 QQKRCQRETERPRHMQIQQRCERYEKEKRRQOQRYEEQOQREDEKYEEMKKEGDNKR 146
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 72 -----GRHQBEEQGRGMHGEEREEHGRGRHGEEREEHGRGRGHN 119
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 147 DQOQREYEDSRNSECQERLQYOCQRRQOQOQROQRG---GDLMNQRGSGRYEEG 202
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 120 GGEEREEGRGHRHGESE-----REERGRGRHGEEREEEGRGRGRGSG 170
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 203 E--EKOSDN--PYUFEDERSLSTRFTEEGHISYLENFYGRSKLLRAKLNRYLLEANDP 258
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 171 ENDEBQDSRRYVVEGRPRFRRIQSDHGFVARALRFQDVSRLKCRITRYVAIMVNR 230
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 259 AVLPYTHLDADAILLVIGRGALKMIRHNDRESYNLECGDVIIRIPAGTTFYLINRDNER 318
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 231 AEVVPGFTDADGVYVAGQEGVLTIVENGEKRSYTKVEGDVIYVAPAGSIMHLANTDGRK 290
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 319 LHAKEFLQITSTPGQYKEFPFAGGONPERPYLSTESKELLEALNTQTERLQVLOQ--- 375
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 291 LVIAKLHTISVPGK-QPL-----SVKPLASLSRYVLRARFKTSDELERLFPNQSQ 344
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 376 ---REGVIIIRASQEOIRLETRDSE---SRHWHIRGSGSSGPYNLFKRPRLYSNKYQ 429
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 345 EKTRASVIVRASEEQLRELRLREAEEQOGHNPRLPRFGSDSDFTLLQDRKIANRHR 404
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 430 AVEVPREDYROLQMDVSVFIANTITOGSMGPFNTSTKVVVVASGEADVEMACPRL-- 487
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 405 LYEADARSPHDLAENDVAVANITPGSMATAYLNTQSKLAVVLEGEDEVQIVCPHLR 464
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 488 ---SGRHGRG-----CGKRHEEEVEVH--EYVRRARLSKREAIIVLAG 526
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 465 ESESEREMKGRREEEEDDQOQRRRGSESESESESESESESESESESESESESESESESE 524
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 527 HPEVYVSS--GNENLLLFAGGINAONNHENFLAGRENVLDQIEPQAMELAPASKEVE 584
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 525 HPEVLEISSQSGSSNQVCFEINAEKNERKWLAGR-NNVIGKISGPAQDLITGRRAREVQ 583
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 585 ELFNQDEST-FFPGPROHOOSPRSTKQOQ 614
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 581 KEVEELFNQDESEFFPPGPRHOQOQSPRSTKQOOPLVSTL 620
DB 503 QAVEELKKNQRESYFVADAPKKEGNKGRK--GPLESSIL 540

RESULT 11

Q01674 PRELIMINARY; PRT; 545 AA.
AC Q01674;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CONVICTILIN PRECURSOR.
OS *Viola namboensis*.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; *Viola*.
OX NCBI_TaxID=3912;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COTYLEDON;
RA Koch G., Koenig S., Becker C., Horstmann C., Schlesiher B.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z71986; CA96513.1; -.
DR HSSP: P50477; ICAU.
DR MENDEL: 12432; *Viola*; 1188; 12432.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
DR PRODOM: PD081059; -; 1.
KW Signal; Seed storage protein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 545 75 GLOBULIN, CONVICTILIN.
SQ SEQUENCE 545 AA; 62810 MW; 459A876F92F5A87E CRC64;

Query Match 24.4%; Score 811; DB 10; Length 545;
Best Local Similarity 34.3%; Pred. No. 1,8e-51;

Matches 196; Conservative 101; Mismatches 204; Indels 70; Gaps 14;

QY 68 ICEEEYENQRPDOOQOYECQCKRCQRETEPRHMOICQRCRRYKREKROKRYEE- 126
DB 23 LCYIYANYDGTSTPRVYQREGRQEGKEKEKH-----GEMRSHKEAOGRRERW 75
QY 127 QQREDEKYEERKKEGDKRDPOQREYEDCRHCEQOEPRLOYQOCRCOQROHGRG 186
DB 76 ETSEEEERVDYEWGRSQRHEDPEERARE--RYRAEERERRRQWE----- 117
QY 187 DLNPNQGGSGRYEEGEEKSD--NPYPDEKSLSTRFTEEGHISYLENFYGRSKLLRA 244
DB 118 -----GEEKEGSSKSQERRNPFLFKSNKFLTLFENENGHIRLQRFDRSLFEN 167
QY 245 LKNYRLVLEENAFVLPHTLDADAILVIGRGALKMHRDNRESYNLECGDVIR 304
DB 168 LQNYRLVEYAKRHTFLPHIDADLLVYLSRALLTVSPDRNSYNLEKEDTILKLP 227
QY 305 GTTFYFLINRDNRLIAKFLQITSTPGQYKEFFPAGQONPEPYLSTFSKEILEALNTQ 364
DB 228 GTTSYLLNODDEDLRYVLDLSISVNRPKVESFGSLGSKN--QYLGFKSNILEASLNTK 285
QY 365 TELRGVY-----GQORE-----GVYIRASQOINELTRDSESRMHIRGCESS 410
DB 286 YETIEKVLLEEDPOOSIGOKRRSQROETNMLVYKSREQVELKRLAKSSS---KKGVSSE 341
QY 411 RGYNLFNKRPLYSNKYGQAYEYKPE--DYROLQDMVSVFIINITQSGMMGPFFNTRSK 469
DB 342 FEFPNLRSONPKYSNKGKLFETPEPKTPOLDLDIFVSSVEINSGMLPHYNRAIV 401
QY 470 VVVASGEADVENACPHLSGRHGGGCGKGRHEEEEVHYEYVARLSKREAIYVLGHV 529
DB 402 ILTVNKGKGLKLE-----LVGKKNQOQORERDEQOV--QRYEARLSPGDVYIIPGHV 454
QY 530 VVPSGGENIILTFAGINAMONNHNELAGERNVLOQIEPQAMELAFASRKEVEELFN 589

DB 455 AVSASSMLNLIL--GFGINAEONRNEIYQSDNVISQIENPVKELTFPGSAQFVNRLKN 512
QY 590 QDESIFFPGPRHOQOQSPRSTKQOOPLVSTL 620
DB 513 QEHSHF--ANAEPEQKEGESQRRKRSPISTL 541

RESULT 12

Q09SP11 PRELIMINARY; PRT; 489 AA.
AC Q09SP11;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SUCROSE BINDING PROTEIN HOMOLOG S-64.
GN SBP.
OS glycine max (soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; *Glycine*.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Pedra J.H.F., Delu-Filho N., Pirovani C.P., Contim L.S., Dewey R.E.,
RA Otoni W.C., Fontes E.P.B.;
RT Antisense and sense expression of a sucrose binding protein homologue
RT gene from soybean in transgenic tobacco affects plant growth and
RT carbohydrate partitioning in leaves.";
RL Plant Sci. 0:0-0(1999).
DR EMBL: AF191299; AAF05723.1; -.
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
SQ SEQUENCE 489 AA; 55834 MW; 9BBC0D45DECECD2 CRC64;

Query Match 23.4%; Score 777.5; DB 10; Length 489;
Best Local Similarity 35.0%; Pred. No. 4,5e-49;

Matches 171; Conservative 99; Mismatches 174; Indels 45; Gaps 11;

QY 138 RMKEGDKRNP-----QOREYED-----GRHCEQOEPRLOYQOCRCOQROH 182
DB 27 KLETEVEEDPELVYTKHOCQOQROYTESDKRYLCQOCMSKQERKQVEEETREKEEK 86
QY 183 GRGGLDMNPQGGSGRYEEGEEKSDNPYFDE--RSLSTRFTEEGHISYLENFYGRSKL 241
DB 87 -----QOEHEEEOQNPYVEEDKDFSTRVETEGSIVLYKKTETSKL 130
QY 242 LRLAKNYRLVLEENAFVLPHTLDADAILVIGRGALKMHRDNRESYNLECGDVIR 301
DB 131 LQGENEFLAILLEARHTEVSPRHPSEVVLNFKRAVLGLVRESETEKITLEPDMIH 190
QY 302 IPAGTFFYFLINRDNRLIAKFLQITSTPGQYKEFFPAGQONPEPYLSTFSKEILEAL 361
DB 191 IPAGTPLYIVNRDENKLLAMLIHPVSTPGKEFFFGGGRDPESVYLAFSMNVLQAL 250
QY 362 NQTERLRLVLOQGRGVYIRASQOIRLTRDSESRMHIRGCESSRGYNLFNKR 421
DB 251 QTPKGLERLFDNQONGSLFKTISRERVALA--PYKKSWMPP--GGE--SKAOFNFTSKRP 306
QY 422 LYSNKYGQAYEYKPEYDRO--LQDMVSVFIANITQSGMMGPFFNTRSTKVVVVASGEADV 480
DB 307 TFSNGYGRLEVGDPDEKMLQRLNMLFTNITQSKSTHNSIATYIALYMDGRGL 366
QY 481 ENACPHLSGRHGGGCGKGRHEEEEVHYEYVARLSKREAIYVLGHVYVSSGNNLL 540
DB 367 QISCPHSSRSDSK-----HDKSSPSYHRISADLKPGWVFVPPGHPVPLASKENLL 420
QY 541 LFAFGINAMONNHNELAGERNVLOQIEPQAMELAFASRKEVEELFNQSDSTIFPPGR 600
DB 421 ITCEVNVADNKKFTFAGKD--NIVSSLDNVAKELAFNPSEYNGVF--ERKESLFFPEEL 478
QY 601 QHOQOSPRS 609

Db 479 PSEERGRRA 487

RESULT 13

049927 PRELIMINARY: PRT: 483 AA.
AC 049927;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE P54 PROTEIN.
GN P54.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RA Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: Y11207; CAAT2090.1; -.
DR HSSP: P50477; ICAU: P155a; 3005; 28319.
DR MENDEL: 28319; P155a; 3005; 28319.
DR INTERPRO: IPR000005; -.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
DR PROSITE: PS00041; HTM_ARAC_FAMILY_1; UNKNOWN_1.
DR PRODOM: PD081059; -. 1.
SQ SEQUENCE 483 AA; 54662 MW; 8127BDAA0178F3D CRC64;

Query Match 23.3%; Score 775.5; DB 10; Length 483;
Best Local Similarity 30.4%; Pred. No. 6, 2e-49;
Matches 170; Conservative 112; Mismatches 163; Indels 115; Gaps 12;

QY 50 QOESDRQOQYCORCKEICEEEYNNRDPQOQYEOCOKRCORRETERHMOICQORC 109
DB 28 RKEKDE-----LTJCKDC-----DMQRYDEDEK-----ICMERC 60
QY 110 ERYEYKRRKQKRYEEQREDEKYEEMKEDNKRDPQOREYEDCRHCEQERLQY 169-
DB 61 D-DYKKKQERQKHEEEDEDEDEDE----- 87
QY 170 OCORRCQEOQROHGRGGLMNPQRGSGRYEEGEKQSDNPYFYDERSLSTRTEGH 229
DB 88 -----ENPYFEEDNDEETKIDTKDGRV 109
QY 230 SVLENFYGSKLRALKNRVLVLEAPNPAFVLPHTLDADAILVIGRGALEKMIHRDN 289
DB 110 LILNKFNEKSKLKIENYGLAVLEIKANAFISPHYDSEALFENIKRGILGLVAEDRT 169
QY 290 ESNVLECGVIRIPAGTTFYLLINRDNERRLHTAKF---LQITSPQOYKEFFAGQNE 346
DB 170 EERNLEEGDMRPVPAATPVLRDENEXLYIAAFMPPSSGSAVLPDEFESAGKRP 229
QY 347 PYLSTESKEILEALNTQTERLGVLGQOREGYIIRASOEQIRELTRDSESRMHIRG 406
DB 230 SVLNTFSKVLQALAKSSNGELETVDQKGRIFIEKEDVGLAPKKS---LWPF--- 283
QY 407 GESSRGPYLNFKNRPLYSNKYGOAYEVKPEDYRQ-LQDMQVSVFANTITQSGMGEFNT 465
DB 284 GGFKSPFNIIFSNNPAFSPKFGSLFEVGPSEKSGLEGILMLTLANTITGSMSTHYNT 343
QY 466 RSTKVVYVSGEADVMACPHLSGRHGRGGCKRHEEEVHEVQVRAKSKREALVYLA 525
DB 344 NANKTALVLDGEBELMACPHPS-----SSNSRQKSSISHTNINAKLRPGVMFVRP 398
QY 526 GHPVYVSSGSENTLLFAGINAOHNHNFVLAGRENNVLOQIEPQAMELAFASRKEVE 585
DB 399 GHPVYVIAKKKKKLLIVVCFEYVNAQNRKKLALAGK-KNITSALDKAKVAFDLAAKVD 457

QY 586 LFNSODESIFFGPROHQO 605
DB 458 VFERKEE-FFPPYDNERKE 476

RESULT 14

022121 PRELIMINARY: PRT: 417 AA.
AC 022121;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE BETA SUBUNIT OF BETA CONGLYCININ.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. WASESUZUNARI;
RA Maruyama N., Katsube T., Mada Y., De La Rosa A., Utsuni S.;
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB008679; BAA23361.1; -.
DR HSSP: P02853; 2PHL.
DR MENDEL: 25075; GLYMA; 1188; 25075.
DR INTERPRO: IPR001113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
DR PRODOM: PD081059; -. 1.
SQ SEQUENCE 417 AA; 48106 MW; DDEE50B5B779556E CRC64;

Query Match 22.8%; Score 758; DB 10; Length 417;
Best Local Similarity 38.0%; Pred. No. 9, 7e-48;
Matches 163; Conservative 87; Mismatches 145; Indels 34; Gaps 7;

QY 205 KOSDNPIYF-DEKSLSTRTEBGIISVLENFYGSKLRALKNRVLVLEAPNPAFVLP 263
DB 6 EDENNPFFRSSSFQTLFENONGRIQLRQFNKRSPOLENLDRVIOFOSKPNITLLP 65
QY 264 THLDADAILLVIGRGALMKMIHRDNRESYNLCGDVIRIPAGTTFYLLINRDNERRLHT 323
DB 66 HHADAFLLFVLSGRAILTLVNNDDRSYNLHPGDQRIAGTYYLVNPHDQNLTK 125
QY 324 FLQITSPQOYKEFFAGQNEPPEYLFSTESKEILEALNTQTERLGVLP-----GQOR 377
DB 126 LAIPVKKPGRYDFFLSNQAOQSYLQGFSHNILETSFSEEEINRVLFGEEDGRQOE 185
QY 378 GYIIRASQEIRELTRDSESRMHIRRGESSRGPNLPFNKRPLYSNKYGOAYEVKPED 437
DB 186 GIVVELSKQIROLISRRAKSSSKRTI---SSDEDFNRSRNPITSYNNFGKFFETPER 241
QY 438 YRLODMQVSVFANTITQSGMGPFFNTSTKVVVVASGEADVEMACPHLSGRHGRGG 497
DB 242 NPQLRDLDFLSSVDNBEALLPNNSKAIYLVINBDANIELV-----G 288
QY 498 -----KRHEEEVHEVQVRAKSKREALVYLAGHPVYVSSGSENTLLFAGINAOHN 551
DB 289 IKRQOQKQKQEEPLEVQVYRAELSEDVFIIPAAYP--FVNANYSNLPFLAFIGIAENN 346
QY 552 HENFLAGRENNVLOQIEPQAMELAFASRKEVEBELFNSODESIFFGPROHQOQSRSTK 611
DB 347 ORNFLAGEKDNVVRQIRVOBELAFPGSAQDVERLLKKORSEFYDA--QPOOKEGSGK 404
QY 612 OQOPLVSTL 620
DB 405 RKQPPSIL 413

RESULT 15
043626 PRELIMINARY: PRT: 438 AA.
ID 043626
AC 043626;

